wherein data division to said first number of sets and calculation of said first total degree of randomness are repeated while a form of data division to said first number of sets is changed, and said group of data is classified into data belonging to the respective classification sets of said first number of classification sets in which said first total degree of randomness is minimized.

59. (New) The method according to claim 58, wherein data division to said first number of sets is performed for data to be classified in numerical order of data values.

60. (New) The method according to claim 58, wherein said calculating the sum of degrees of randomness in the respective sets of said first number of sets comprises:

estimating a probability distribution of data values in each of said sets on the basis of said data values of said data belonging to each of said sets;

obtaining an entropy of each of said estimated probability distributions of data values; and

weighting said entropy of each of said probability distributions in accordance with the number of data belonging to a corresponding one of said sets.

61. (New) The method according to claim 60, wherein said first probability distribution is a normal distribution.

62. (New) The method according to claim 58, further comprising:

dividing data belonging to a specific classification set in said first number of classification sets into a second number of sets having no common elements; and

calculating a second total degree of randomness which is a sum of degrees of randomness of data values in the respective sets of said second number of sets,

wherein data division to said second number of sets and calculation of said second total degree of randomness are repeated while a form of data division to said second number

of sets is changed, and said data belonging to said specific classification set are further classified into data belonging to the respective classification sets of said second number of classification sets in which said second total degree of randomness is minimized.

63. (New) The method according to claim 62, wherein data division to said second number of sets is performed for data to be classified in numerical order of data values.

64. (New) The method according to claim 62, wherein said calculating the sum of degrees of randomness in the respective sets of said second number of sets comprises:

estimating a probability distribution of data values in each of the sets on the basis of said data values of said data belonging to each of said sets;

obtaining an entropy of each of the estimated probability distributions of data values; and

weighting said entropy of each of said probability distributions in accordance with the number of data belonging to a corresponding one of said sets.

65. (New) The method according to claim 64, wherein said first probability distribution is a normal distribution.

66. (New) A data classification apparatus for classifying a group of data into a plurality of sets in accordance with data values, comprising:

a first data dividing unit which divides said group of data into a first number of sets having no common elements; and

a first degree-of-randomness calculation unit which calculates degrees of randomness of data values in the respective sets divided by said first data dividing unit, and calculates a sum of the degrees of randomness; and

a first classification unit which is electrically connected to the first degree-of-randomness calculation unit and classifies said group of data into said data belonging to the

respective classification sets of said first number of classification sets in which said sum of degrees of randomness calculated by said first degree-of-randomness calculation unit is minimum out of forms of data division by said first data dividing unit.

68. (New) The apparatus according to claim 66, further comprising:

a second data dividing unit which divides data belonging to a specific classification set in the first number of classification sets into a second number of sets having no common elements; and

a second degree-of-randomness calculation unit which is electrically connected to the second data dividing and calculates degrees of randomness of data values in the respective sets divided by said second data dividing unit and calculates a sum of the degrees of randomness; and

a second classification unit which is electrically connected to the second degree-of-randomness calculation unit and classifies said data of said specific classification set into said data belonging to the respective classification sets of said second number of classification sets in which said sum of degrees of randomness calculated by said second degree-of-randomness calculation unit is minimum out of forms of data division by said second data dividing unit.

69. (New) A signal processing method of processing a measurement signal obtained by measuring an object, comprising:

extracting signal levels at a plurality of feature points obtained from said measurement signal; and

setting said extracted signal levels as classification object data and classifying said signal levels at said group of feature points into a plurality of sets by using the data classification method according to claim 58.

70.
69. (New) The method according to claim 68, wherein said feature point is at least one of a local maximum point and a local minimum point of said measurement signal.

71.
70. (New) The method according to claim 68, wherein said feature point is a point of inflection of said measurement signal.

A1 Cont'd

72.
71. (New) A signal processing apparatus for processing a measurement signal obtained by measuring an object, comprising:

a measurement unit which measures said object and acquires a measurement signal;

an extraction unit which is electrically connected to the measurement unit and extracts signal levels at a plurality of feature points obtained from said measurement signal; and

the data classification apparatus according to claim 66, which sets said extracted signal levels as classification object data.

73.
72. (New) A position detection method of detecting a position of a mark formed on an object, comprising:

acquiring an image pick-up signal by picking up an image of said mark;

processing said image pick-up signal as a measurement signal by said signal processing method according to claim 68; and

calculating said position of said mark on the basis of a signal processing result obtained in said signal processing.

74.
73. (New) The method according to claim 72, wherein

in data classification in said signal processing, the number of data which should belong to each classification set after said data classification is known in advance, and

in said position calculation, the number of data which should belong to each classification set is compared with the number of data in each of said classification sets

classified in said signal processing to evaluate validity of the classification in said signal processing, and said position is calculated on the basis of said data belonging to said classification set evaluated to be valid.

74. (New) A position detection apparatus for detecting a position of a mark formed on an object, comprising:

an image pick-up unit which acquires an image pick-up signal by picking up an image of said mark;

the signal processing apparatus according to claim 71, which performs signal processing for said image pick-up signal as a measurement signal; and

a position calculation unit which calculates said position of said mark on the basis of a signal processing result obtained by said signal processing apparatus.

75. (New) An exposure method of transferring a predetermined pattern onto a divided area on a substrate, comprising:

detecting a position of a position detection mark formed on said substrate by the position detection method according to claim 72, obtaining a predetermined number of parameters associated with a position of said divided area, and calculating arrangement information of said divided area on said substrate; and

transferring said pattern onto said divided area while performing position control on said substrate on the basis of said arrangement information of said divided area obtained in said arrangement calculation.

76. (New) An exposure apparatus for transferring a predetermined pattern onto a divided area on a substrate, comprising:

a substrate stage on which said substrate is mounted; and

the position detection apparatus according to claim 74, which detects a position of said mark on said substrate.

78. (New) A data classification method of classifying a group of data into a plurality of sets in accordance with data values, comprising:

classifying said group of data into a first number of sets in accordance with said data values; and

dividing said group of data again into a second number of sets which is smaller than said first number on the basis of a characteristic of each of said first number of sets divided in data classification into said first number of sets.

79. (New) The method according to claim 77, wherein data classification into said second number of sets comprises:

specifying a first set, of said first number of sets, which meets a predetermined condition;

estimating a first boundary candidate for dividing said group of data excluding data included in said first set by using a predetermined estimation technique;

estimating a second boundary candidate for dividing a data group, of said group of data, which is divided by said first boundary candidate and includes said first set by using said predetermined estimation technique; and

dividing said group of data into said second number of sets on the basis of said second boundary candidate.

80. The method according to claim 78, wherein said predetermined estimation technique comprises:

calculating a degree of randomness of data values in each set divided by said boundary candidate, and calculating a sum of said degrees of randomness; and

performing said degree-of-randomness calculation while changing a form of data division with said boundary candidate, and extracting a boundary candidate with which said sum of degrees of randomness obtained in said degree-of-randomness calculation is minimized.

80. (New) The method according to claim 78, wherein said predetermined estimation technique comprises:

obtaining a probability distribution in each set of said data group; and

extracting said boundary candidate on the basis of a point of intersection of said probability distributions of the respective sets.

81. (New) The method according to claim 78, wherein said predetermined estimation technique comprises:

calculating an inter-class variance as a variance between sets divided by said boundary candidate; and

performing said intra-class variance calculation while changing a form of data division with said boundary candidate, and extracting a boundary candidate with which the inter-class variance obtained in said inter-class variance calculation is maximized.

82. (New) The method according to claim 78, wherein said predetermined condition is a condition that data exhibiting a value substantially equal to a predetermined value is extracted from said group of data.

83. (New) The method according to claim 82, wherein

said group of data is image pick-up data of the respective pixels obtained by picking up different image patterns within a predetermined image pick-up field; and

said predetermined value is image pick-up data of a pixel existing in an area corresponding to an image pick-up area for a predetermined image pattern.

84. The method according to claim 77, wherein said dividing data into said second number of sets comprises:

extracting a predetermined number of sets from the first number of sets on the basis of the number of data included in the respective sets of said first number of sets;

calculating an average data value by averaging data values respectively representing sets of said predetermined number of sets; and

dividing said group of data into said second number of sets on the basis of said average data value.

85. The method according to claim 84, wherein in said average data value calculation, a weighted average of said data values is calculated by using a weight corresponding to at least one of the number of data of the respective sets of said predetermined number of sets and a probability distribution of said predetermined number of sets.

86. The method according to claim 77, wherein said group of data is luminance data of the respective pixels obtained by picking up different image patterns within a predetermined image pick-up field.

87. (New) A data classification apparatus for classifying a group of data into a plurality of sets in accordance with data values, comprising:

a first data dividing unit which divides said group of data into a first number of sets on the basis of said data values; and

a second data dividing unit which is electrically connected to the first data dividing unit and divides said group of data into a second number of sets smaller than said first number again on the basis of a characteristic of each of said first number of sets.

89.

88. (New) The method according to claim 87, wherein said first number is not less than three, and said second number is two.

90.

89. (New) An image processing method of processing image data obtained by picking up an image in a predetermined image pick-up field, comprising:

setting luminance data, as a group of data, which is obtained by picking up an image pattern of an object and an image pattern of a background which exist in said predetermined image pick-up field; and

identifying a boundary between said object and said background by classifying said luminance data by using the data classification method according to claim 86.

91.

90. (New) The method according to claim 89, wherein said object includes a substrate onto which a predetermined pattern is transferred.

92.

91. (New) An image processing apparatus for processing image data obtained by picking up an image in a predetermined image pick-up field, wherein

luminance data, which is obtained by picking up an image pattern of an object and an image pattern of a background which exist in said predetermined image pick-up field is set as a group of data, and

a boundary between said object and said background is identified by classifying said luminance data by using the data classification apparatus according to claim 87.

93.

92. (New) An exposure method of transferring a predetermined pattern onto a substrate, comprising:

specifying an outer shape of said substrate by using the image processing method according to claim 90;

controlling a rotational position of said substrate on the basis of said specified outer shape of said substrate;

detecting a mark formed on said substrate after said rotational position is controlled; and

transferring said predetermined pattern onto said substrate while positioning said substrate on the basis of a mark detection result obtained in said mark detection.

94.

~~93~~. (New) An exposure apparatus for transferring a predetermined pattern onto a substrate, comprising:

an outer shape specifying unit including the image processing apparatus according to claim 91, which specifies an outer shape of said substrate;

a rotational position control unit which is electrically connected to the outer shape specifying unit and controls a rotational position of said substrate on the basis of said outer shape of said substrate which is specified by said image processing apparatus;

a mark detection unit which detects a mark formed on said substrate whose rotational position is controlled by said rotational position control unit; and

a positioning unit which is electrically connected to the mark detection unit and positions said substrate on the basis of a mark detection result obtained by said mark position detection unit,

wherein said predetermined pattern is transferred onto said substrate while said substrate is positioned by said positioning unit.

95.

~~94~~. (New) A data classification method of classifying a group of data into a plurality of sets in accordance with data values, comprising:

estimating a first number of boundary candidates for dividing said group of data into a second number of sets on the basis of said data values; and

extracting a third number of boundary candidates which is smaller than said first number and is used to divide said group of data into a fourth number of sets smaller than said

second number, under a predetermined extraction condition, on the basis of said first number of boundary candidates.

96.

95. (New) The method according to claim 94, wherein said predetermined extraction condition includes a condition that said third number of boundary candidates are extracted on the basis of a magnitude of a data value indicated by each of said first number of boundary candidates.

97.

96. (New) The method according to claim 95, wherein said predetermined extraction condition includes a condition that a boundary candidate with which said data value is maximized is extracted.

98.

97. (New) The method according to claim 94, wherein

said group of data are arranged at positions in a predetermined direction, and

said predetermined extraction condition includes a condition that said fourth number of boundary candidates are extracted on the basis of the respective positions of said first number of boundary candidates.

99.

98. (New) The method according to claim 94, wherein

said group of data are differential data obtained by differentiating image pick-up data of the respective pixels obtained by picking up different image patterns in a predetermined image pick-up field in accordance with positions of said pixels,

said data value is a differential value of said image pick-up data, and

said boundary candidate is a position of said pixel.

100.

99. (New) The method according to claim 94, wherein said first number is not less than two, and said third number is one.

101.

100. (New) The method according to claim 94, wherein said group of data are luminance data of the respective pixels obtained by picking up different image patterns in a predetermined image pick-up field.

102.

101. (New) A data classification apparatus for classifying a group of data into a plurality of sets in accordance with data values, comprising:

a first data dividing unit which estimates a first number of boundary candidates for dividing said group of data into a second number of sets on the basis of said data values; and

a second data dividing unit which is electrically connected to the first data dividing unit and extracts a third number of boundary candidates which is smaller than said first number and is used to divide said group of data into a fourth number of sets smaller than said second number, under a predetermined extraction condition, on the basis of said first number of boundary candidates.

103.

102. (New) The apparatus according to claim 101, wherein

said group of data are differential data obtained by differentiating image pick-up data of the respective pixels obtained by picking up different image patterns in a predetermined image pick-up field in accordance with positions of said pixels,

said data value is a differential value of said image pick-up data, and

said boundary candidate is a position of said pixel.

104.

103. (New) The apparatus according to claim 101, wherein said first number is not less than two, and said third number is one.

105.

104. (New) An image processing method of processing image data obtained by picking up an image in a predetermined image pick-up field, comprising:

setting luminance data, as a group of data, which is obtained by picking up an image pattern of an object and an image pattern of a background which exist in the predetermined image pick-up field; and

identifying a boundary between said object and said background by classifying said luminance data by using the data classification method according to claim 94.

105. (New) An image processing apparatus for processing image data obtained by picking up an image in a predetermined image pick-up field, wherein

luminance data which is obtained by picking up an image pattern of an object and an image pattern of a background which exist in said predetermined image pick-up field is set as a group of data, and

a boundary between said object and said background is identified by classifying said luminance data by using the data classification apparatus according to claim 101.

106. (New) An exposure method of transferring a predetermined pattern onto a substrate, comprising:

specifying an outer shape of said substrate by using the image processing method according to claim 104;

controlling a rotational position of said substrate on the basis of said specified outer shape of said substrate;

detecting a mark formed on said substrate after said rotational position is controlled; and

transferring said predetermined pattern onto said substrate while positioning said substrate on the basis of a mark detection result obtained in said mark detection.

107. (New) An exposure apparatus for transferring a predetermined pattern onto a substrate, comprising:

an outer shape specifying unit including the image processing apparatus according to claim 105, which specifies an outer shape of said substrate;

a rotational position control unit which controls a rotational position of said substrate on the basis of said outer shape of said substrate which is specified by said image processing apparatus;

a mark detection unit which detects a mark formed on said substrate whose rotational position is controlled by said rotational position control unit; and

a positioning unit which positions said substrate on the basis of a mark detection result obtained by said mark position detection unit,

wherein said predetermined pattern is transferred onto said substrate while said substrate is positioned by said positioning unit.

108. (New) A recording medium on which a position detection control program executed by a position detection apparatus for detecting a position of a mark formed on an object is recorded, wherein

said position detection control program comprises:

allowing an image of said mark to be picked up and allowing an image pick-up signal to be acquired;

a signal processing control program using said image pick-up signal as a measurement signal, comprising

allowing signal levels at a plurality of feature points obtained from said measurement signal to be extracted; and

said data classification control program using said extracted signal levels as a group of classification object data, comprising

allowing said group of data to be divided into a first number of sets having no common elements;

allowing a first total degree of randomness which is a sum of degrees of randomness of data values in the respective sets of said first number of sets to be calculated; and

allowing said group of data to be divided into data belonging to the respective classification sets of said first number of classification sets in which said first total degree of randomness is minimized, by repeating data division to said first number and calculation of said first total degree of randomness while changing a mode of data division to said first number of sets; and

allowing a position of said mark to be calculated on the basis of a processing result on said image pick-up signal.

116.

~~109~~. (New) The medium according to claim 108, wherein

in said data classifying, the number of data which should belong to each classification set after said data classification is known in advance, and

the number of data which should belong to each classification set is compared with the number of data in each of said classified classification sets to evaluate validity of said data classifying, and said position is calculated on the basis of data belonging to said classification set evaluated to be valid.

111.

~~110~~. (New) A recording medium on which an image processing control program executed by an image processing apparatus for processing image data obtained by picking up an image in a predetermined image pick-up field is recorded, wherein

said image processing control program comprises:

allowing luminance data, which is obtained by picking up an image pattern of an object and an image pattern of a background which exist in said predetermined image pick-up field, to be set as a group of data;

a data classification control program which allows said luminance data to be classified, comprising:

allowing said group of data to be divided into a first number of sets on the basis of said data values; and

allowing said group of data to be divided into a second number of sets smaller than said first number again on the basis of features of the respective first number of sets; and

allowing a boundary between said object and said background to be identified.

11. (New) A recording medium on which an image processing control program executed by an image processing apparatus for processing image data obtained by picking up an image in a predetermined image pick-up field is recorded, wherein

said image processing control program comprises:

allowing luminance data which is obtained by picking up an image pattern of an object and an image pattern of a background which exist in said predetermined image pick-up field to be set as a group of data;

a data classification control program which allows said luminance data to be classified, comprising

allowing a first number of boundary candidates for dividing said group of data into a second number of sets to be estimated on the basis of said data values;

allowing a third number of boundary candidates which is smaller than said first number and is used to divide said group of data into a fourth number of sets smaller than said